NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

ANALYSIS OF PREDICTIVE FACTORS FOR FULLY MISSION CAPABLE RATES OF DEPLOYED AIRCRAFT

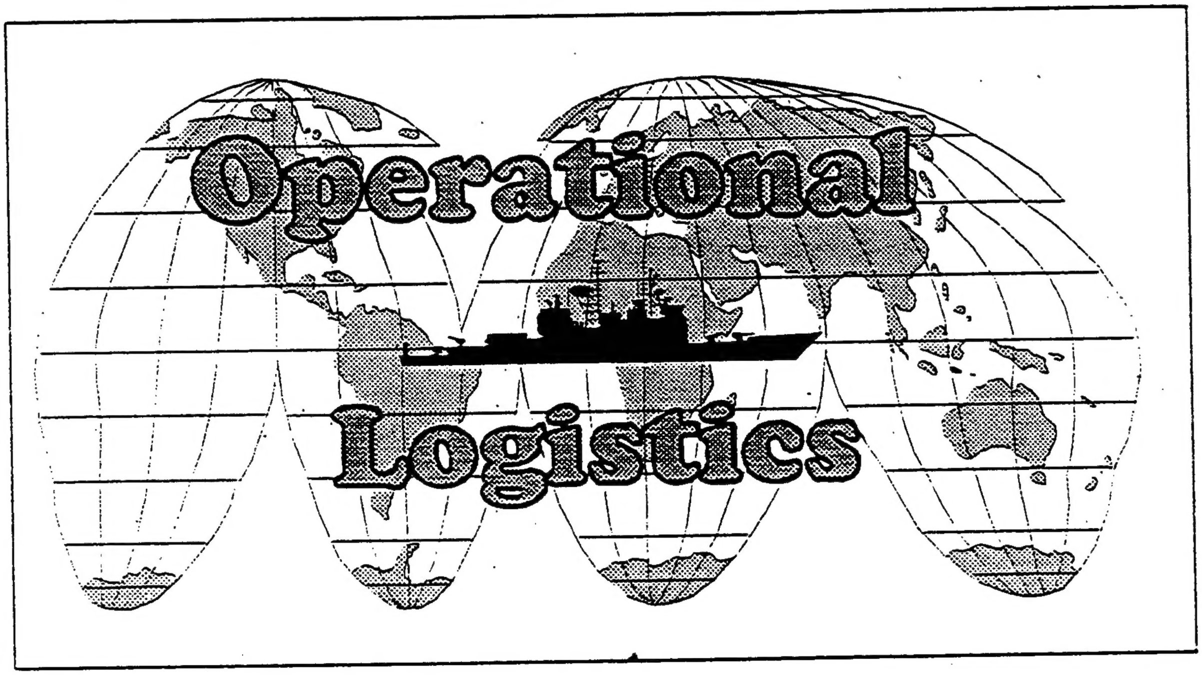
by

Patricia B. Moore

September 1998

Thesis Advisor: Samuel E. Buttrey
Second Reader: Lyn R. Whitaker

Approved for public release; distribution is unlimited.


Operational
Logistics

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 1998	Master's Thesis

4. TITLE AND SUBTITLE
ANALYSIS OF PREDICTIVE FACTORS FOR FULLY MISSION CAPABLE RATES OF DEPLOYED AIRCRFT

5. FUNDING NUMBERS

6. AUTHOR(S)
Moore, Patricia B.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

As the U.S. military reduces its forces, the ability to maintain an acceptable level of readiness is of concern to the U.S. Navy. Both personnel and equipment readiness and the ability to predict them have been the focus of much attention. Fully Mission Capable (FMC) rates measure the percentage of time that aircraft are fully able to meet the mission requirements. FMC rates have been determined to be the best single measure of equipment condition, providing an indication of aircraft readiness. This thesis evaluates the capabilities of logistic regression and regression trees in predicting aircraft readiness for a specific carrier deployment or aircraft type/model/series (TMS). The data are taken from observations of squadrons by aircraft TMS by month from 1981 through 1997. Empirical results indicate that logistic regression and regression trees provide better forecasting results in predicting aircraft or airwing readiness than taking the mean and standard deviation of the historical data.

14. SUBJECT TERMS
Aircraft Readiness, Fully Mission Capable Rates, Logistic Support, Logistic Regression, Regression Trees

15. NUMBER OF PAGES
67

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

Approved for public release; distribution is unlimited

ANALYSIS OF PREDICTIVE FACTORS FOR FULLY MISSION CAPABLE RATES OF DEPLOYED AIRCRFT

Patricia B. Moore
Lieutenant Commander, United States Navy
B.A., State University of New York, Binghamton, 1982

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1998

Author: ______________________
Patricia B. Moore

Approved by: ______________________
Samuel E. Buttrey, Thesis Advisor

____
Lyn R. Whitaker, Second Reader

____
Richard E. Rosenthal, Chairman
Department of Operations Research

ABSTRACT

As the U.S. military reduces its forces, the ability to maintain an acceptable level of readiness is of concern to the U.S. Navy. Both personnel and equipment readiness and the ability to predict them have been the focus of much attention. Fully Mission Capable (FMC) rates measure the percentage of time that aircraft are fully able to meet mission requirements. FMC rates have been determined to be the best single measure of equipment condition, providing an indication of aircraft readiness. This thesis evaluates the capabilities of logistic regression and regression trees in predicting aircraft readiness for a specific carrier deployment or aircraft type/model/series (TMS). The data are taken from observations of squadrons by aircraft TMS by month from 1981 through 1997. Empirical results indicate that logistic regression and regression trees provide forecasting results with standard errors of prediction better than taking the mean and standard deviation of the historical data.

TABLE OF CONTENTS

LIST OF ACRONYMS

AAW	Anti-air Warfare
AFQT	Armed Forces Qualification Test
AID	Aircraft Information Database
AIMD	Aviation Intermediate Maintenance Department
AVCAL	Aviation Consolidated Allowance List
AVDLR	Aviation Depot Level Repairable
ASW	Anti-submarine Warfare
BCM	Beyond Capability Maintenance
CASREP	Casualty Report
CCC	Command Control and Communication
CINC	Commander in Chief
CNA	Center for Naval Analyses
CNO	Chief of Naval Operations
CVBG	Carrier Battle Group
EMR	Enlisted Master Record
EW	Electronic Warfare
FHP	Flying Hour Program
FLE	Fatigue Life Extension
FOUO	For Official Use Only
FMC	Fully Mission Capable
GLM	Generalized Linear Model
HQMC	Headquarters Marine Corps
MC	Mission Capable
MTBF	Mean Time Between Failures
MTW	Major Theater War
NAVAIR	Naval Air
NAVICP	Navy Inventory Control Point
NAVSUP	Naval Supply
NMC	Not Mission Capable
NMCM	Not Mission Capable Maintenance
NMCS	Not Mission Capable Supply
NOFORN	No Foreigners
OMR	Officer Master Record
OPNAV	Naval Operations
PMCM	Partial Mission Capable Maintenance
PMCS	Partial Mission Capable Supply
RBS	Readiness Based Sparing
RFI	Ready for Issue
SORTS	Status of Resource and Training System
STW	Strike Warfare
TMS	Type/Model/Series
TRANSCOM	Transportation Command
TYCOM	Type Commander
VA	Attack Squadron

VF	Fighter Squadron
VFA	Strike Squadron

EXECUTIVE SUMMARY

As the U.S. military reduces its forces, the ability to maintain an acceptable level of readiness is of great concern to the U.S. Navy. Both personnel and equipment readiness and the ability to predict them have been the focus of much attention. The Department of Defense Dictionary of Military and Associated Terms defines readiness as "the ability of forces, units, weapons systems, or equipments to deliver the output for which they are assigned." Aircraft readiness is defined as the percentage of time the aircraft in a squadron are able to perform all of their missions such as anti-air warfare (AAW), anti-submarine warfare (ASW), strike warfare (STW), command control and communications (CCC) and electronic warfare (EW). Fully Mission Capable (FMC) rates measure the percentage of time that aircraft are fully able to meet mission requirements and provide an indication of aircraft readiness. In a study conducted by the Center for Naval Analyses (CNA), FMC rates were found to provide the best summary measure of aircraft equipment condition.

The Chief of Naval Operations (CNO) provides FMC goals for deployed aircraft by type/model/series (TMS) as well as an overall FMC goal for all aircraft. The current goal for the overall FMC rate for deployed aircraft is 61%. The FMC rates goals for those aircraft TMS having the greatest negative impact on deployed FMC rates are 55% for the F-14A, 55% for the F-14B, 66% for the F-14D and 63% for the F/A-18C. While overall FMC rates have been relatively stable since Fiscal Year 1991 (FY91), FMC rates for deployed aircraft fell sharply in FY97. The average FMC rate in 1997 for all reporting aircraft on the four deployed carriers was 58.9%. The average FMC rate during the previous six years, which included data from 31 deployments, had been approximately

69.6% with no deployed carrier reporting a deployed FMC rate of less than 59.9%. The USS Constellation reported an FMC rate of 52.8% and the USS Independence reported and FMC rate of 44.1% during their 1997 deployments.

The intent of this thesis is to build models that identify the most important FMC rate predictors and interactions for specific aircraft TMS as well as for deployed airwings. This thesis addresses several specific hypotheses: 1) the greater the percentage of requests for repairable and consumable items which take longer than 2 days to fill, the lower the FMC rate; 2) the higher the cannibalization rates, the lower the FMC rates; 3) as flight hours increase, FMC rates decrease; 4) the quantity and quality of personnel directly affects FMC rates; and 5) a model developed using all observations can be used to predict FMC rates for a particular aircraft TMS or airwing deployed on a particular aircraft carrier.

Given the above hypotheses, this analysis also has the research goals of identifying significant predictive factors (not found in previous research), comparing and contrasting the logit regression and the regression tree methodologies for this type of data set, and addressing the implication of the resulting predictive models.

Logistic (logit) regression and regression trees are used to model the relationship between FMC rates and groups of variables that measure: 1) aircraft usage such as flight hours and flight hours between failures; 2) maintenance practices such as cannibalization rates, percentages of items requiring depot-level maintenance and average time to respond to a request for a repairable; and 3) the quality and quantity of personnel such as the number of enlisted people assigned to the squadron, the percentage of the crew who were not with the squadron three months earlier, and the percentage of the crew with a

high school diploma. Both the logit model and the regression tree result in estimates of aircraft readiness. These are compared against CNO's goal and to test the predictive power of the models, the data set will be randomly divided into two sets without replacement. The first set is used to build models and the second set is used to test their predictive capabilities. In addition, the predictive power of the models is tested on small subsets of data for a specific aircraft carrier deployment or aircraft TMS.

There are three main implications of these predictive models that result from the data analysis. The first is that the models can predict FMC rates with some success and are an improvement over having no model at all. The second implication is that AVCAL variables are significant factors in predicting FMC rates. The third implication is that in addition to the models' predictive power, resource managers are provided with a list of significant predictive factors on which to focus time and money in an effort to improve aircraft readiness.

I. INTRODUCTION

A. OVERVIEW

As the U.S. military reduces its forces, the ability to maintain an acceptable level of readiness is of great concern to the U.S. Navy. Both personnel and equipment readiness and the ability to predict them have been the focus of much attention. [Refs. 5, 6, 9, 10] The Department of Defense Dictionary of Military and Associated Terms defines readiness as "the ability of forces, units, weapons systems, or equipments to deliver the output for which they are assigned." [Ref. 8] Aircraft readiness is defined as the percentage of time the aircraft in a squadron are able to perform all of their mission areas. Mission areas include anti-air warfare (AAW), anti-submarine warfare (ASW), strike warfare (STW), command control and communications (CCC) and electronic warfare (EW). In a study conducted by the Center for Naval Analyses (CNA), Fully Mission Capable (FMC) rates were found to provide the best summary measure of aircraft equipment condition. [Ref. 9]

B. BACKGROUND

FMC rates measure the percentage of time that aircraft are fully able to meet mission requirements and were determined to be the best single measure of equipment condition, providing an indication of aircraft readiness [Ref. 9]. CNA came to this conclusion by using principal components to compute linear combinations or weighted averages of variables measuring material condition. These original variables were equipment Status of Readiness and Training System (SORTS), FMC rates, mission capable (MC) rates, partially mission capable due to supply (PMCS) rates, partially mission capable due to maintenance (PMCM) rates, not mission capable (NMC) rates,

not mission capable due to supply (NMCS) rates, and not mission capable due to maintenance (NMCM) rates. Principal components compute as many linear combinations as original variables. Each linear combination "explains" successively less of the available information. The first linear combination of the material condition variables computed, explained about 76% of the total variability of these variables. In fact, the first linear combination had more than a 99% correlation with the fully mission capable rate. It was concluded that the additional linear combinations provided no advantage and that therefore the FMC rate was the best measure of aircraft readiness. [Ref. 9] The Chief of Naval Operations (CNO) provides FMC goals for deployed aircraft by type/model/series (TMS) as well as an overall FMC goal for all aircraft. The current goal for the overall FMC rate for deployed aircraft is 61%. CNA determined that the top four aircraft having the greatest negative effect on deployed FMC rates were the F-14A, F-14B, F-14D and the F/A-18C. The FMC rates goals are 55% for the F-14A, 55% for the F-14B, 66% for the F-14D and 63% for the F/A-18C. [Ref. 10]

While overall FMC rates have been relatively stable since Fiscal Year 1991 (FY91), FMC rates for deployed aircraft fell sharply in FY97. The average FMC rate in 1997 for all reporting aircraft on the four deployed carriers was 58.9%. The average FMC rate during the previous six years, which included data from 31 deployments, had been approximately 69.6% with no deployed carrier reporting a deployed FMC rate of less than 59.9%. In 1997, the USS Constellation reported a deployed FMC rate of 52.8% and the USS Independence reported a deployed FMC rate of 44.1%, both lower than the established goal. The Commander, Theodore Roosevelt Battle Group, raised concerns about the low FMC rates during a post-deployment brief following the Group's

deployment from December 1996 to May 1997 [Ref. 5]. As a result of these concerns, a Task Force was formed under the leadership of the Assistant Deputy Chief of Naval Operations (Logistics) that included representatives from Naval Operations (OPNAV) N4/N8, Headquarters Marine Corps (HQMC), Fleet Commanders in Chief (CINCs)/Air Type Commanders (TYCOMs), Naval Air (NAVAIR), Naval Supply (NAVSUP), Naval Inventory Control Point (NAVICP), CNA, and the Naval Center for Cost Analysis. Their mission was to evaluate whether spare parts problems were becoming more prevalent, just-in-time readiness was forcing more cannibalization because of delayed transportation, Navy surged multi-Carrier Battle Groups (CVBGs) were causing shortages by competing for existing parts and transportation, or inventories were eroding.

The task force analyzed trends in over 150 readiness, maintenance, transportation, wholesale supply support, and retail supply support metrics. Their review confirmed that both maintenance and supply factors influenced material support for Naval aviation. The Task Force concluded that three significant influences explained the drop in the FMC rate during Fiscal Year 1997. First, aircraft deployed in a lower material condition because of insufficient funding for stocking consumables at air stations in support of squadron and Aviation Intermediate Maintenance Department (AIMD) maintenance requirements. Second, the F-14A/B/D and the F/A-18C experienced several unusual maintenance problems that led to a decrease in the mean flying hours between failures and an increase in maintenance downtime. And third, actual Flying Hour Program costs exceeded the budgeted level by about 20%, which resulted in a cascading impact on spares availability.

C. PROBLEM DEFINITION

The intent of this thesis is to build models that identify the most important FMC rate predictors and variable interactions for specific aircraft TMS as well as for deployed airwings. This thesis will address several specific hypotheses: 1) the greater the percentage of requests for repairable and consumable items which take longer than 2 days to fill, the lower the FMC rate; 2) the higher the cannibalization rates the lower the FMC rates; 3) as flight hours increase, FMC rates decrease; 4) the quantity and quality of personnel directly effects FMC rates; and 5) a model developed using all observations, can be used to predict FMC rates for a particular aircraft TMS or airwing deployed on a particular aircraft carrier.

Given the above hypotheses, this analysis also has the research goals of identifying significant predictive factors (not found in previous research), comparing and contrasting the logit regression and the regression tree methodologies for this type of data set, and addressing the implication of the resulting predictive models.

Logistic (logit) regression and regression trees will be used to model the relationship between FMC rates and groups of variables that measure: 1) aircraft usage such as flight hours and flight hours between failures, 2) maintenance practices such as cannibalization rates, percentages of items requiring depot-level maintenance and average time to respond to a request for a repairable, and 3) the quality and quantity of personnel such as the number of enlisted people assigned to the squadron, the percentage of the crew who were not with the squadron three months earlier, and the percentage of the crew with a high school diploma. Both the logistic regression and regression tree models predict the FMC rates by values bounded by zero and one. Predicted FMC rates from

these models will be compared to CNO's goals to see whether aircraft readiness is acceptable under various conditions. To test the predictive power of the models, the data set used in this thesis will be randomly divided into two sets without replacement. The first set will be used to build models and the second set will be used to test the predictive capabilities of the models.

In the next chapter, the current literature on aircraft readiness is reviewed. Chapter III discusses the methodology behind logistic regression and regression trees. Chapters IV and V provide the results and conclusions of the analysis.

II. LITERATURE REVIEW

CNA has conducted several studies on readiness in recent years. One study of particular interest looked at several measures of personnel, training, and equipment readiness [Ref. 9]. The expectation was that FMC determinants could be obtained from among the following categories: 1) the quality or quantity of personnel; 2) the wear on the equipment; 3) the availability of spare parts; and 4) the design of the aircraft. The FMC rate was selected as the best measure of equipment readiness and was estimated using two methods.

The first method fit a regression model to FMC rates using data from 1982 through 1994: Results of this analysis were then used to predict the 1995 data (as if the 1995 data had not been collected). The regression model used fighter and attack squadron FMC rates as the dependent variable with the explanatory variables being TMS indicators, enlisted turnover, enlisted personnel quality, sorties, deployed sorties, months since last deployed, current deployment status, the equipment readiness of the carrier, last month's FMC rate, and supplies on hand. Several conclusions were drawn from that analysis. First, the quality of the enlisted squadron personnel was strongly linked to FMC rates. The second conclusion was that as the squadron moved farther in time from its last deployment, it spent more days per month in a FMC ready status. Thirdly, deployed sorties were associated with lower FMC rates, and finally, the TMS of aircraft helped explain a squadron's FMC rate. A major problem with this analysis was the inability of the regression model to predict 1995 FMC rates. There were large errors between the actual and predicted FMC rates.

The second method used in [Ref. 9] was a Markov chain model. The model was constructed to look at the probability of an aircraft moving from a fully mission capable status to a not fully mission capable status and the probability of moving back. The model realistically accounted for and reported the FMC rates, which were represented as an aggregation of information about individual aircraft, and for the autocorrelation between observations at adjacent times. The development of the model was not completed, but initial analysis of subsets of the data, consisting of F/A-18s and F-14s, suggested that the Markov chain model held promise. However, there were still unresolved methodological issues associated with numerical estimation algorithms, goodness-of-fit measures, and hypothesis tests. In general, they found that measures of personnel readiness, supply, and the number of sorties were strongly associated with FMC rates [Ref. 9]. Using similar techniques, a master's thesis has developed a Markov chain model to model readiness for F/A-18 aircraft on the USS Independence during its most recent 1998 deployment. [Ref. 1] For this application, predicting FMC rates using a Markov chain model was at least as good as, if not better than, those obtained using classical regression techniques.

In more recent studies, CNA looked at managing readiness as well as the volatility in readiness measures [Refs. 6, 10]. Budget cuts have required a closer look at where funds are allocated. The concern is that the wrong allocation of scarce resources will result in a sudden and unrecoverable decline in readiness. CNA determined that there are three approaches to readiness. The first is to determine that a future decrease is likely and act now to prevent or mitigate it. The second is to establish a healthy long-term environment so that adverse changes are less likely to occur; the third is to wait until

the change occurs and then act to stop and reverse the change. The second approach, which deals with preventing a decline in readiness, is the focus of this thesis. In order to manage readiness and prevent such a decline, the key drivers of aircraft readiness must be identified through statistical models. [Ref. 10] CNA's research, however, indicates that because of the variance in the data and the volatility in the FMC rates which vary considerably from month to month, it is difficult to forecast readiness accurately more than a few months into the future. [Ref. 6]

III. METHODOLOGY

A. DESCRIPTION OF DATA

CNA maintains an extensive aircraft database, containing data for 94 different fighter (VF), attack (VA), and strike (VFA) squadrons from April 1980 to December 1995. The database, classified secret, was updated in June and July 1998 to include all carrier-based aircraft from August 1980 through April 1998. There were a total of 32,894 records and 49 variables in the data set sorted by aircraft TMS and squadron in several computer spreadsheets. The data sources were the SORTS database, the Aircraft Information Database (AID), the Enlisted Master Record (EMR), the Aviation Consolidated Allowance List (AVCAL) database, the Naval Sea Logistics Center and the ship employment history database.

The data include a set of variables that 1) measure aspects of a squadron's deployment cycle; 2) measure the frequency at which aircraft were flown; 3) describe personnel quality and quantity; and finally, 4) measure supply-related issues.

The data was consolidated into one spreadsheet and imported in to S-Plus (Mathsoft Inc., 1997) for the analysis. The data set includes various measures of aircraft equipment condition and readiness such as FMC rates, MC rates, PMCS rates, PMCM rates, NMC rates, NMCS rates, NMCM rates and equipment SORTS. Any one of these measures can be used as the dependent variable in a model. However, based on prior research by CNA, the FMC rates are selected as the best measure of aircraft readiness and are used in this analysis.

Initially, it was thought that the SORTS scoring system would be a good indicator for predicting FMC rates. SORTS is a type of grading system for the status of resources

in the areas of personnel, supply, equipment, and training. It focuses on the sufficiency of resources and the completion of training requirements. To determine the SORTS score, each squadron computes a series of ratios in each area considered, and measures the quantity of the resource on hand compared with the authorized requirement. The ratios are then transformed into an ordinal score with five levels, C1, ..., C5, with C1 indicating the highest level of readiness. The SORTS scoring system set provided an indication of readiness problems within the areas of personnel, supply, equipment, and training but did not explain the causes behind the score [Ref. 8]. Because the scores could not be explained, it was felt that it would be difficult to explain any type of predictive power the data might have. As a result, the entire SORTS data, which was the classified portion of the data set, was removed from the data set. The data set was reclassified as No Foreigners (NOFORN) and For Official Use Only (FOUO).

Prior to beginning the analysis, the data was checked for missing information. Null fields and fields containing only "." were replaced with NA. Closer examination of the data set reveal that all records for the E-2C, EA-6B, ES-3A, and S-3B are missing AVCAL variables that detail repairable and consumable items onboard the carriers. In addition, the E-2C is missing enlisted personnel data from the EMR. Because it is difficult to fit logistic regression models with missing data, some models are fit excluding AVCAL variables and where possible, models are fit using all available variables.

Because the primary objective was to predict FMC rates for carrier-based aircraft, records for aircraft from 10 reserve squadrons and 15 training squadrons were removed from the data set. Next, aircraft that were no longer in the Navy's inventory or no longer deployed onboard carriers were removed from the data set. Once these records and

variables were removed, 12,397 records and 38 variables remained in the data set. Table 1 gives the 38 variables used in the analysis.

The dependent variable *fmc*, marked in Table 1 with "**", is a number between 0 and 100 and measures the percent of time that aircraft were able to meet the missions that they were required to meet. Because it is meant to be an accurate portrayal of the availability of mission ready aircraft, it includes downtime associated with all maintenance actions. The independent variables available for FMC analysis are marked in Table 1 with a "*". The character variables used to describe the record such as *linkuic, acft, squadron,* and *depflg* were used to subdivide the data set for more detailed analysis of a particular aircraft carrier or deployment cycle.

Table 1: Data Description

Variable	Description
uic	Unit Identification Code for squadron
year	Year of the current observation
month	Month of the current observation
acft	Type/Model/Series (TMS) of aircraft
squadron	Name of the squadron
linkuic	UIC of the base/ship the squadron is attached
primac	Binary (0,1), 1 if aircraft TMS is primary one for squadron-month
depflg	Binary (0,1), 1 if squadron is deployed
numac	Number of aircraft in squadron/month/TMS
msind	Number of months since last deployment
fmc**	Fully mission capable rate
util*	Utilization rate of aircraft: flight hours/number of aircraft
repr*	Percent of items processed at AIMD which were repaired
bcm*	Percent of items processed at AIMD which were Beyond the Capability of Maintenance
	and were sent to depot-level maintenance
aimd*	Number of items processed at the AIMD
canns*	Number of cannibalizations for the squadron in current month per 100 flight hours
dmmh*	Direct maintenance man hours per flight hour
fh*	Flight hours for squadron
etma*	Elapsed maintenance time per maintenance action
fhbf*	Flight hours between failures
fhbma*	Flight hours between maintenance actions
sorties*	Number of sorties flown by squadron
numppl*	Number of enlisted personnel assigned to the squadron
crewhsdg*	Percent of crew with high school degree
smart*	Percent of crew who scored in upper mental group on AFQT
smarter*	Percent of crew who scored in AFQT Categories I and II
cfstprom*	Percent of crew who make E5 in 4 years or less
demote*	Percent of crew who had a higher paygrade last quarter
clos*	Average length of service in months
turn3mo*	Percent of crew who were not with the squadron 3 months earlier
turn6mo*	Percent of crew who were not with the squadron 6 months earlier
wgtmann*	Percent of crew who are attached to the squadron compared with M+1 requirements
	weighted by paygrade
rp01*	Percent of requests for repairable items that were filled in 1 or 2 days
rrst*	Average time it took to respond to a request for a repairable item
rreq*	Number of requests for repairable items
cp01*	Percent of requests for consumable items that were filled in 1 or 2 days
crst*	Average time it took to respond to a request for a consumable item
creq*	Number of requests for consumable items

B. DATA SUBSETS FOR ANALYSIS AND TESTING

Two data subsets were used to develop two separate models. Removing all records with missing AVCAL values, leaves a data set with 5,858 records and 38 variables (including AVCAL variables concerning repairable and consumable items). The independent variables from the AVCAL data source are *rp01, rrst, rreq, cp01, crst,* and *creq,* and are defined in Table 1. This data set was used to build one model for the F-14A, F-14B, F-14C, F/A-18A and the F/A-18C. A second model was built on the data set excluding AVCAL variables; this data set contains 10,923 records and 32 variables for analysis. By omitting the columns with the AVCAL data, the model was built from observations of all aircraft TMS still remaining in the data set but with six fewer independent variables. Before each model was built, the data was randomly divided into two parts. Then first part with 66% of the data was used to build the models, with the remaining 33% saved to test their predictive powers.

In addition, the models were tested separately on five subsets of the data: an east coast carrier, the USS George Washington, during its 1996 deployment; a west coast carrier, the USS Carl Vinson, during its 1996 deployment; a forward deployed carrier based out of Yokosuka, Japan, the USS Independence, during 1997; and finally the F-14A and the F/A-18C across all deployments.

1. Descriptive Statistics of Data Set without AVCAL Variables

Of the 10,923 records, 3,146 observations were from F-14As, 1,929 from E-2Cs, 1,803 from EA-6Bs, 1,680 from F/A-18Cs, 944 from F/A-18As, and 925 from S-3Bs; the remaining 496 were from F-14Bs and F-14Ds. The dates of the observations were from 1981 through 1997. The mean FMC rate was 60.29%. The mean percent of items

processed at the AIMD which were beyond the capability of local maintenance and were sent to depot-level maintenance was 29.21%. The mean number of cannibalizations for the squadron in the current month per 100 flight hours was 29.15. The mean number of sorties was 140.7. The mean percentage of the crew with high school degrees was 88.4% and the mean percent of crew who scored in the upper mental group on the AFQT was 54.92%.

2. Descriptive Statistics of Data Set with AVCAL Variables

Of the 5,858 records, 2,994 observations were from F-14As, 1,531 from F/A-18Cs, 919 from F/A-18As, 269 from F-14Bs and 145 from F-14Ds. The dates of the observations were from 1982 through 1996. The mean FMC rate was 62.06%. The mean percent of items processed at the AIMD which were beyond the capability of local maintenance and were sent to depot-level maintenance was 26.98%. The mean number of cannibalizations for the squadron in the current month per 100 flight hours was 29.42. The mean number of sorties was 190.4, which is quite a bit higher than the mean number of sorties from the data set without AVCAL variables. The mean percentage of the crew with high school degrees was 88.24%. The mean percentage of crew who scored in the upper mental group on the AFQT was 53.58%. From the AVCAL variables, the mean percentage of requests for repairable items that were filled in 1 or 2 days was 74.09%, the mean percentage of requests for consumable items that were filled in 1 or 2 days was 67.26% and the mean number of requests for consumable items was 268.2.

C. LOGISTIC (LOGIT) REGRESSION

1. Explanation of Method

A logistic regression model will be used to predict the percentage of the time an aircraft or airwing deployed onboard a carrier is fully able to perform all of its mission areas, given the values of the independent variables. Logistic regression models are appropriate because the response variable is a rate bounded by 0 and 1. The predicted FMC rates will be compared with CNO's goals for the overall readiness of an airwing or readiness for a specific aircraft TMS. For example, for an airwing, percentages greater than or equal to CNO's goal of 61% represent the "ready" status while percentages less than 60% represent the "not ready" status. The FMC rate either meets the target CNO goal or it does not. The dependent variable, *fmc*, is derived from an observation of an aircraft TMS in a particular squadron by month and year.

As with linear regression, the logistic regression model uses maximum likelihood estimation to estimate the probability of obtaining the observed response as a function of the independent variables. However, a linear regression model can give probability estimates with values greater than 1 or less than 0. The logistic regression model is more realistic in predicting "ready" or "not ready" FMC rates because the model requires that all predictions fall within the 0-1 range. In this model, the observed response, *fmc* (expressed here as a percentage), will be divided by 100 to obtain a number between 0 and 1.

A brief description of the model used to predict the percentage of time an aircraft in a squadron or airwing is able to meet all mission requirements is given below. Let Y_i

$i = 1,2,\ldots$ be a binary variable that takes value 1 if the aircraft is ready (FMC) and 0 otherwise, then:

$Y_i \sim$ Bernouilli (p_i), $\quad i = 1, 2, \ldots$

$\log(p_i / (1-p_i)) = \beta_0 + X_i\beta,$

where

p_i = the probability the aircraft is FMC for observation i, $\quad i = 1, 2, \ldots$

β_0 = a coefficient to be estimated,

X_i = a row vector of K–1 explanatory variables that may influence the FMC rate for observation i, $\quad i = 1, 2, \ldots$

β = a column vector of K–1 regression coefficients to be estimated.

Based on the logit model, the probability that $Y_i = 1$ can also be defined as:

$$p_i = \frac{1}{(1 + e^{-L_i})}, \tag{1}$$

where

$$L_i = \beta_0 + \sum_{k=1}^{K-1} \beta_k X_{ik}, \quad i = 1, 2\ldots \tag{2}$$

and X_{ik}, $i = 1, 2, \ldots, k = 1, 2, \ldots$ K–1, is the value of the k^{th} explanatory variables for observation i. When $X_{i1} = X_{i2} = \ldots = X_{iK-1} = 0$, the probability $P(Y_i = 1)$ is $p = 1/(1 - e^{-\beta_0})$. Each one-unit increase in X_{ik}, for a specific explanatory variable, increases the odds favoring $Y_i = 1$ by $100(e^{\beta_k} - 1)$ percent. [Ref. 6]

2. Procedures

Each data frame was constructed with the response variable, *fmc,* and the explanatory variables in order to use logistic regression in S-Plus. Because there were so many independent variables, backward selection was used initially to remove variables

which did not improve the basic model with the response variable modeled by all explanatory variables. The method of model selection is based on Akaike's Information Criterion (*AIC*) statistic that provides a convenient criterion for determining whether the model is improved by dropping a term. [Ref. 13] When no term's deletion will lower the *AIC*, the process stops. In the case of the model without the AVCAL variables, the model began with 21 explanatory variables and ended with 14 explanatory variables. In the case of the model with the AVCAL variables, the model began with 27 explanatory variables and ended with 18 explanatory variables. The next step in the analysis was to use stepwise model selection of 2nd order interactions. The goal was to produce a statistically sound model with predictive powers using the fewest explanatory variables.

D. REGRESSION TREES

1. Explanation of Method

Regression trees are an alternative to logistic models for uncovering the structure of data. They have been increasingly used to devise prediction rules that can be rapidly and repeatedly evaluated in summarizing large multivariate data sets. [Ref. 2] The difference between a regression tree and a classification tree is that a regression tree's terminal node contains a numeric value, while a classification tree's contains a factor. S-Plus automatically recognizes the type of tree being grown by whether the response variable is numeric (as in the FMC case). While a tree is grown only on those observations without missing values, it offers the advantage of treating missing values more satisfactorily when the model is used to predict an outcome. An observation is dropped down the tree until a missing value is encountered or a terminal node is reached. [Refs. 2, 13]

The process begins with a root or parent node. Tree construction uses a computationally intensive algorithm that searches over all the variables to find the optimal binary split, which produces the maximum reduction in the deviance. That split produces two "child" nodes. The algorithm recursively splits the data in each node until the resulting node either is homogeneous or contains too few observations. The default in S-Plus is five or fewer observations. The deviance of the tree is defined as the sum of the deviances in the leaf nodes; in node *i* the deviance is

$$D_i = \Sigma_k \, (y_k - \mu_i)^2,$$

where μ_i is the average in node *i* and y_k is the value of the independent variable for observation *k,* counting only observations in node *i.* Each pair of child nodes has a combined deviance that is smaller than its parent node.

In the case of continuous variables, the possible splits depend on the data representation. For example, the *canns* variable is the number of cannibalizations for the squadron per month per 100 flight hours and is tracked with a precision in tenths. So these numbers might be 15.0, 15.1, 15.2 and so on. The node will be split between the values, e.g., above 15.15 and below 15.15.

The tree construction process often results in an extremely large tree that can be too elaborate and over-fit the data. To correct this, cross-validation can be used to determine the optimal size of the tree while pruning enables the analyst to choose the number of terminal leaves, reducing it to a reasonable size. In cross-validation, trees are grown from mutually exclusive subsets of the original data set. Trees are grown on all but one of the subsets; the remaining is used to test the predictive power of the tree. The tree is then pruned until a balance between predictive power and fit is achieved. The

process is repeated until the best size has been determined by finding the minimum deviance across all iterations. [Ref. 15] The resulting optimal tree becomes the model.

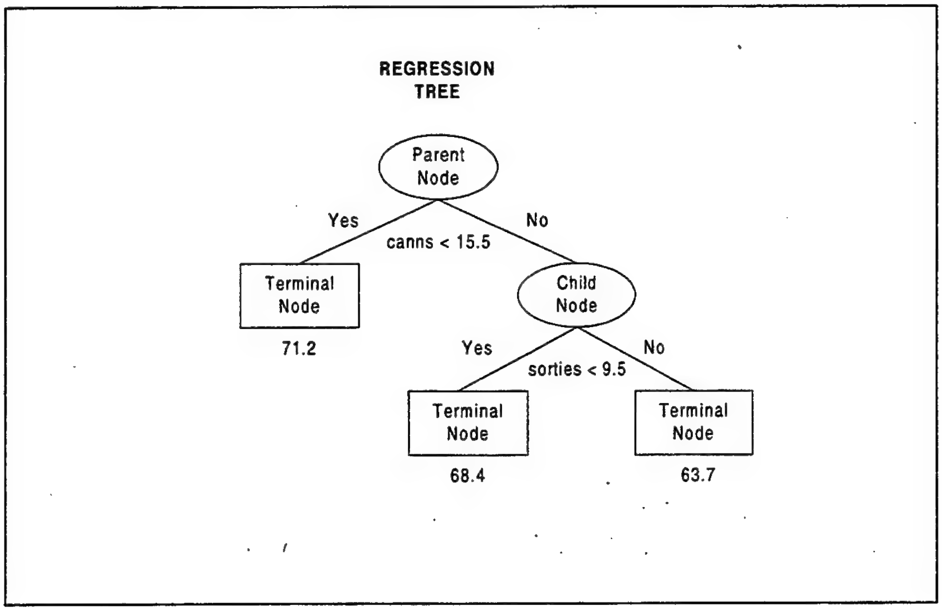


Figure 1: A Hypothetical Tree

Interpreting the results of a tree involves dropping an observation down to a terminal node. For example, Figure 1 would indicate that an airwing with more than 15.5 cannibalizations in the current month per 100 flight hours and fewer than 9.5 sorties per month would predict an FMC rate of 68.4%. At the end of the process the prediction tree structure is easy to understand, interpret, and use.

2. Procedures

A tree model for each data set was built with the response variable, *fmc*, and the explanatory variables using regression tree modeling in S-Plus. The initial tree had 136 leaves and was too large to interpret. Cross-validation identified an optimal tree with 23 leaves for the data set without AVCAL data and with 26 leaves for the set with AVCAL

data. The trees were then pruned to these recommended sizes. The pruning process actually optimizes the deviance penalized by model size. Figures 2 and 3 show the relationship between the deviance and model size measured by the number of leaves in the model.

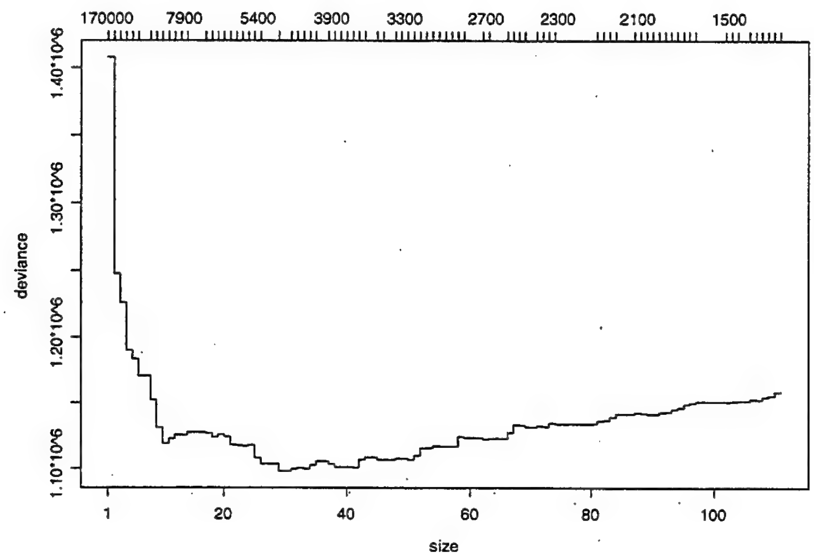


Figure 2: Tree Size versus Deviance (No AVCAL Variables)

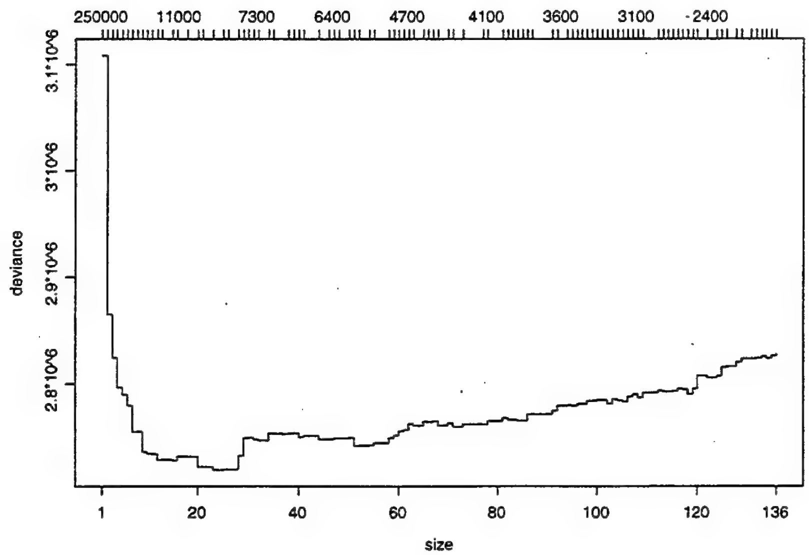


Figure 3: Tree Size versus Deviance (AVCAL Variables)

IV. RESULTS

A. LOGISTIC REGRESSION MODEL (NO AVCAL VARIABLES)

The logistic model for FMC rates using the data frame without AVCAL variables is summarized in Table 2. We call this the full model. The significance of the model can be determined by comparing the difference between the null deviance and the residual deviance to a χ^2 with eleven degrees of freedom. A χ^2 with eleven degrees of freedom has a critical value of 19.68 at a 5% level of significance, which is much smaller than 205.43, the difference between the null deviance and the residual deviance. However, one point should be noted. In a logistic regression model the residual deviance should be approximately equal to the degrees of freedom. In this model, though, the residual deviance is much smaller than the number of degrees of freedom indicating that the data is under-dispersed. That is, the FMC rates do not vary as much as expected, indicating some correlation structure for which the model has not accounted. One possible explanation is that as the FMC rate falls below CNO's goal, a squadron or airwing will make every attempt to raise the rate above the target rate. Once above the rate, a squadron or airwing has little incentive to get their rate even higher. The overall effect is to keep FMC rates fairly constant across all squadrons.

Table 2: Logistic Regression Summary (No AVCAL Variables)

Variables	Est.Value	Std. Error	t-value
(Intercept)	5.1389	1.4700	3.4958
bcm	–0.0202	1.4700	–6.8614
canns	–0.0005	0.0003	–2.1016
sorties	0.0012	0.0004	3.0123
numppl	–0.0257	0.0078	–3.3070
smart	–0.0573	0.0263	–2.1779
canns:sorties	–0.0001	0.0000	–8.1660
dmmh:sorties	0.0001	0.0000	3.5100
etma:cfstprom	–0.0297	0.0070	–4.2395
numppl:cfstprom	0.0005	0.0002	3.3654
bcm:wgtmann	0.0001	0.0000	4.4701
numppl:smart	0.0003	0.0001	2.5440
Null Deviance: 1412.984 on 7360 degrees of freedom Residual Deviance: 1207.545 on 7349 degrees of freedom			

The factors that significantly (5% level of significance) increase FMC rates with an increase in their value are *sorties*, the interactions between *dmmh* and *sorties*, *numppl* and *cfstprom*, *bcm* and *wgtmann*, and *numppl* and *smart*. The factors that significantly decrease FMC rates with an increase in their value, are *bcm*, *canns*, *numppl*, *smart*, and the interactions between *canns* and *sorties*, and *etma* and *cfstprom*. All other variables listed in Table 1 were removed during the backward and stepwise selection procedures. As is often the case in regression models of this type, the signs of the coefficients make the model difficult to interpret. For example, the model appears to indicate that as the number of people assigned to the squadron who scored in the upper mental group on the AFQT increases, the FMC rate decreases. To facilitate the use of a model as a decision-making tool, a reduced model was developed eliminating variables that were too complex or not easily explained. This reduced model is summarized in Table 3.

Table 3: Logistic Regression Summary (No AVCAL Variables)

Variables	Est. Value	Std. Error	t-value
(Intercept)	0.7577	0.0894	8.4731
bcm	–0.0136	0.0024	–5.6964
sorties	0.0025	0.0004	8.3869
canns:sorties	–0.0001	0.0000	–10.3049
Null Deviance: 1412.984 on 7360 degrees of freedom Residual Deviance: 1269.406 on 7357 degrees of freedom			

The significant predictors in this model are much easier to explain. The factors which significantly decrease FMC rates with an increase in their value, are *bcm* and the interaction between *canns* and *sorties*. The factor that increases FMC rates with an increase in its value, is *sorties*. While not as good as the model with eleven parameters, a χ^2 with three degrees of freedom has a critical value of 7.815 at a 5% level of significance, which is much smaller than 143.57, the difference between the null deviance and the residual deviance. And once again, the model indicates that the data is under-dispersed.

B. REGRESSION TREE MODEL (NO AVCAL VARIABLES)

Cross-validation identified as optimal a tree with 23 terminal nodes. The tree model is depicted in Figure 4. The root node indicates a predicted FMC rate of 60.31%, which is the average over the entire data set. The number inside each node is the predicted FMC rate while the number below the node is the number of observations in that node. The rectangular nodes are terminal nodes or leaves for the tree. The first split divides the data into two sets: those observations with fewer than 15.15 cannibalizations for the squadron in the current month per 100 flight hours and those with more than 15.15 cannibalizations. For example: a deployed carrier with fewer than 7.55 cannibalizations per squadron per 100 flight hours, more than 31 sorties, and less than 56.55% of items

processed at the AIMD classified "Beyond the Capability of Maintenance" and sent to depot-level maintenance would have an FMC rate of 71.4%. Of the 7,361 observations in the test data set, 1,347 are classified along this path. One interesting observation of the tree is that 52.6% of the observations fall into just four leaves. A second is that the range of the predicted FMC rates in the model is from 7% to 71.37% while the range of the actual FMC rates from the historical data is from 0% to 100%. A third observation is that 47% of the observations fall down the left half of the tree into 8 leaves and that 53% of the observations fall down the right half of the tree into 15 leaves. This indicates that although the data splits roughly in half at the root node, observations with fewer than 15.15 cannibalizations become homogenous much more quickly.

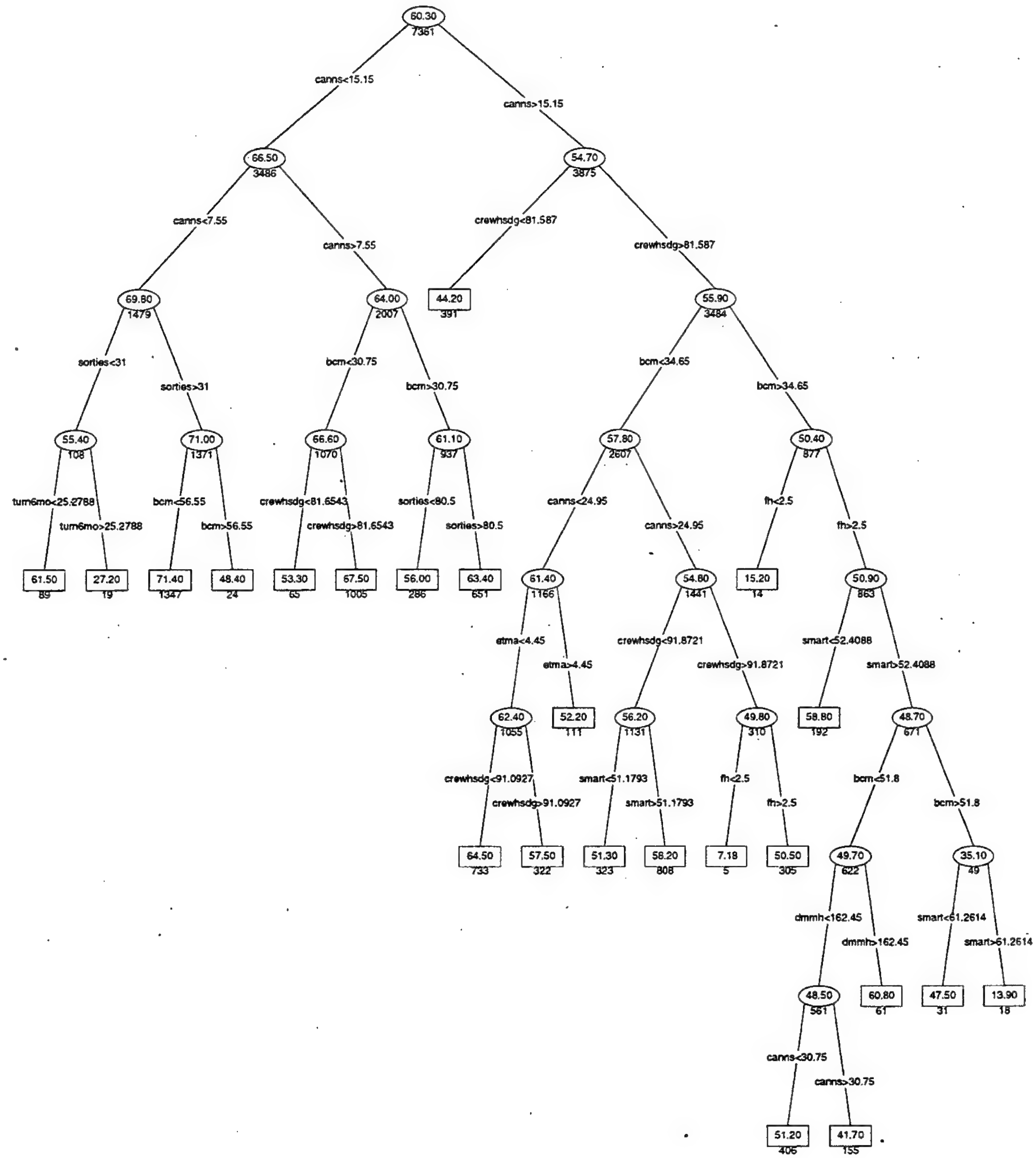


Figure 4: Tree Model Without AVCAL Variables

C. PREDICTIONS

A summary of the predicted results is depicted in Table 4. The test set is the 33% of the data that was held out for testing the predictive powers of the models. The remaining subsets were built from the original data set and may include observations used to build the models. The table lists the mean and standard deviation of the subsets' historical data as well as the predicted mean and the standard error of prediction from the full logistic regression model, the reduced logistic regression model and the regression tree model. With the exception of the USS George Washington's 1996 deployment, the models predict better than just taking the mean and standard deviation of the historical data. The USS George Washington's average FMC rate of 80.82% is well above the average FMC rate of 60.29% of the data set used to build the model. The models' inability to accurately predict the USS George Washington's FMC rate could be because the ship's data is so different from the rest of the data. In the post-deployment brief there is an indication that the aviation support personnel addressed potential problems that could negatively effect the FMC rate early on and took the necessary steps to prevent a degradation of FMC rates. During the pre-deployment phase a comprehensive and complete AVCAL was developed and a detailed review of demand data from previous deployments was conducted in order to ensure that the essential parts were on board prior to the deployment. The models do not adequately account for the intangibles that superior performers bring to the equation. In all cases considered, the regression tree model is the best predictor of FMC rates with the smallest standard error of prediction.

When the models are used to predict whether or not the deployment or aircraft TMS will meet CNO's FMC rate goals (i.e., ready or not ready), the models predict well

for all subsets except for the USS Independence. For the USS Independence, which is considered in a deployed status at all times even when in port in Yokosuka Japan, the historical data has an FMC rate of 61.89%, just above CNO's goal, while each model predicts just below the goal.

Table 4: Data without AVCAL Variables

SUBSET	MEAN	STD ERR
Test Set (33% of Data)		
Historical	60.26	20.37
Full Model	60.46	20.34
Reduced Model	60.32	20.34
Regression Tree	60.14	18.89
USS George Washington		
Historical	80.82	9.96
Full Model	65.03	9.95
Reduced Model	63.78	9.96
Regression Tree	63.98	9.66
USS Carl Vinson		
Historical	63.82	18.44
Full Model	63.97	18.38
Reduced Model	62.73	18.40
Regression Tree	63.98	16.18
USS Independence		
Historical	61.89	22.51
Full Model	58.38	22.34
Reduced Model	59.83	22.47
Regression Tree	58.59	21.05
F-14A		
Historical	61.38	18.60
Full Model	61.38	18.56
Reduced Model	61.68	18.58
Regression Tree	59.41	16.29
F/A-18C		
Historical	68.42	15.00
Full Model	66.48	14.98
Reduced Model	64.51	14.98
Regression Tree	65.54	14.34

As a final check of the goodness of fit of each model, the predictive ability of the models is tested on the remaining 3,562 observations, accounting for 33% of the sample.

Of these 3,562 observations, 1,871 are actually "ready". Each observation is classified in relation to its predicted probability compared with a threshold value of 61%. If the probability is greater than or equal to 61%, it is classified as "ready". If it is less than 61%, it is classified as "not ready". There are two types of errors. A model can incorrectly predict that an aircraft TMS is ready, when it is not (Type II error), and that an aircraft TMS is not ready, when in fact it is (Type I error). These type I and type II errors are considered equally costly when predicting results. The naive model assumes that all aircraft TMS are ready and has an error rate of 47.48%. Table 5 provides a summary of the prediction results.

Table 5: Model Prediction Summary (No AVCAL Variables)

Model	Threshold Probability	Correct Predictions (out of 3562)	Errors	Gain	% Gain over Naive Prediction
Naive		1871	1691		
Full Model	**0.61**	**2267**	**1295**	**396**	**21.17**
Reduced Model	0.61	2195	1367	294	17.32
Regression Tree	0.61	2243	1319	342	19.88

The column "errors" shows the total type I and type II errors. The full logistic regression model, shown in bold, provides the greatest improvement in the percent gain in correct predictions with 21.17%.

D. LOGISTIC REGRESSION MODEL (AVCAL VARIABLES)

The logistic model for FMC rates using the data frame with AVCAL variables is summarized in Table 6. The significance of the model can be determined by comparing the difference between the null deviance and the residual deviance to a χ^2 with twelve degrees of freedom. A χ^2 with twelve degrees of freedom has a critical value of 21.03 at

a 5% level of significance, which is significantly smaller than 125.53, the difference between the null deviance and the residual deviance. Once again, the residual deviance is much smaller than the number of degrees of freedom indicating that the data is under-dispersed.

Table 6: Logistic Regression Summary (AVCAL Variables)

Variables	Est. Value	Std. Error	t-value
(Intercept)	0.9048	1.5906	0.5688
bcm	–0.0193	0.0082	–2.3421
etma	–1.3459	0.3385	–3.9758
sorties	0.0018	0.0004	4.1291
cfstprom	–0.9385	0.2419	–3.8797
wgtmann	0.0307	0.0134	2.2822
cp01	0.0498	0.0180	2.7563
creq	–0.0004	0.0002	–2.1500
canns:sorties	–0.0001	0.0000	–5.8857
wgtmann:cp01	–0.0005	0.0002	–2.5543
crewhsdg:cfstprom	0.0110	0.0028	3.9189
etma:smart	0.0206	0.0057	3.5972
crewhsdg:smart	–0.0006	0.0002	–2.9037
Null Deviance: 641.7771 on 3960 degrees of freedom			
Residual Deviance: 516.2463 on 3947 degrees of freedom			

The factors that significantly (5% level of significance) increase FMC rates with an increase in their values are *sorties*, *wgtmann*, *cp01*, the interactions *crewhsdg* and *cfstprom*, *bcm* and *wgtmann*, and *etma* and *smart*. The factors that significantly decrease FMC rates with an increase in their values, are *bcm*, *etma*, *cfstprom*, *creq* and the interactions between *canns* and *sorties*, *wgtmann* and *cp01*, and *crewhsdg* and *smart*. All other variables that were initially listed in Table 1 were removed during the backward and stepwise selection procedures. As with the logistic regression model built without the AVCAL data, this model is difficult to interpret because of the interactions between personnel variables and aviation maintenance variables. To facilitate the use of the model as a decision-making tool, a reduced model was developed eliminating variables

that were too complex and difficult to explain. This reduced model is summarized in Table 7.

Table 7: Logistic Regression Summary (AVCAL Variables)

Variables	Est. Value	Std. Error	t-value
(Intercept)	0.3704	0.1655	2.2382
etma	–0.0626	0.0308	–2.0313
sorties	0.0018	0.0004	4.3423
cp01	0.0055	0.0015	3.7499
creq	–0.0005	0.0002	–2.4467
canns:sorties	–0.0001	0.0000	–6.0971
Null Deviance: 639.787 on 3949 degrees of freedom			
Residual Deviance: 545.428 on 3944 degrees of freedom			

The tradeoff in selecting a reduced model which is easier to interpret, is that it does not predict as well as the model with twelve parameters. In this case, the factors that significantly decrease FMC rates with an increase in their value, are *etma*, *creq* and the interaction between *canns* and *sorties*. The factors that increase FMC rates with an increase in their values, are *sorties* and *cp01*. It is worth noting that for both models the AVCAL variables, *cp01* and *creq*, are considered significant predictors of FMC rates. While not as good as the model with twelve parameters, a χ^2 with five degrees of freedom has a critical value of 11.07 at a 5% level of significance, which is much smaller than 94.36, the difference between the null deviance and the residual deviance. As was the case in all the previous logistic regression models, the residual deviance is much smaller than the number of degrees of freedom indicating that the data is under-dispersed.

E. REGRESSION TREE MODEL (AVCAL VARIABLES)

Cross-validation identified as optimal a tree with 26 terminal nodes. The tree model is depicted in Figure 5. The root node indicates a predicted FMC rate of 61.97%. As before, the number inside each node is the predicted FMC rate while the number

below the node is the number of observations in that node. The rectangular nodes are terminal nodes or leaves for the tree. The first split divides the data into two sets: those observations with fewer than 19.65 cannibalizations for the squadron in the current month per 100 flight hours and those with more than 19.65 cannibalizations. The deviance is reduced by approximately 11% at the initial split and is reduced at each additional split. For example: a deployed carrier with more than 19.65 cannibalizations per squadron per 100 flight hours, fewer than 87.715% of the requests for repairable items filled in 1 or 2 days, more than 5.5 flight hours per squadron and fewer than 233.5 requests for consumable items would have a predicted FMC rate of 56.9%. One interesting observation of the tree is that 56.8% of the observations fall into just five leaves. A second observation is that the range of the FMC rate is from 17.53% to 75.94% while the range of FMC rates from the actual data is from 0% to 100%. A third observation is that 61% of the observations fall down the left half of the tree into 11 leaves and that 39% of the observations fall down the right half of the tree into 15 leaves. This indicates that although the tree appears fairly well balanced, significantly more observations have fewer than 19.65 cannibalizations.

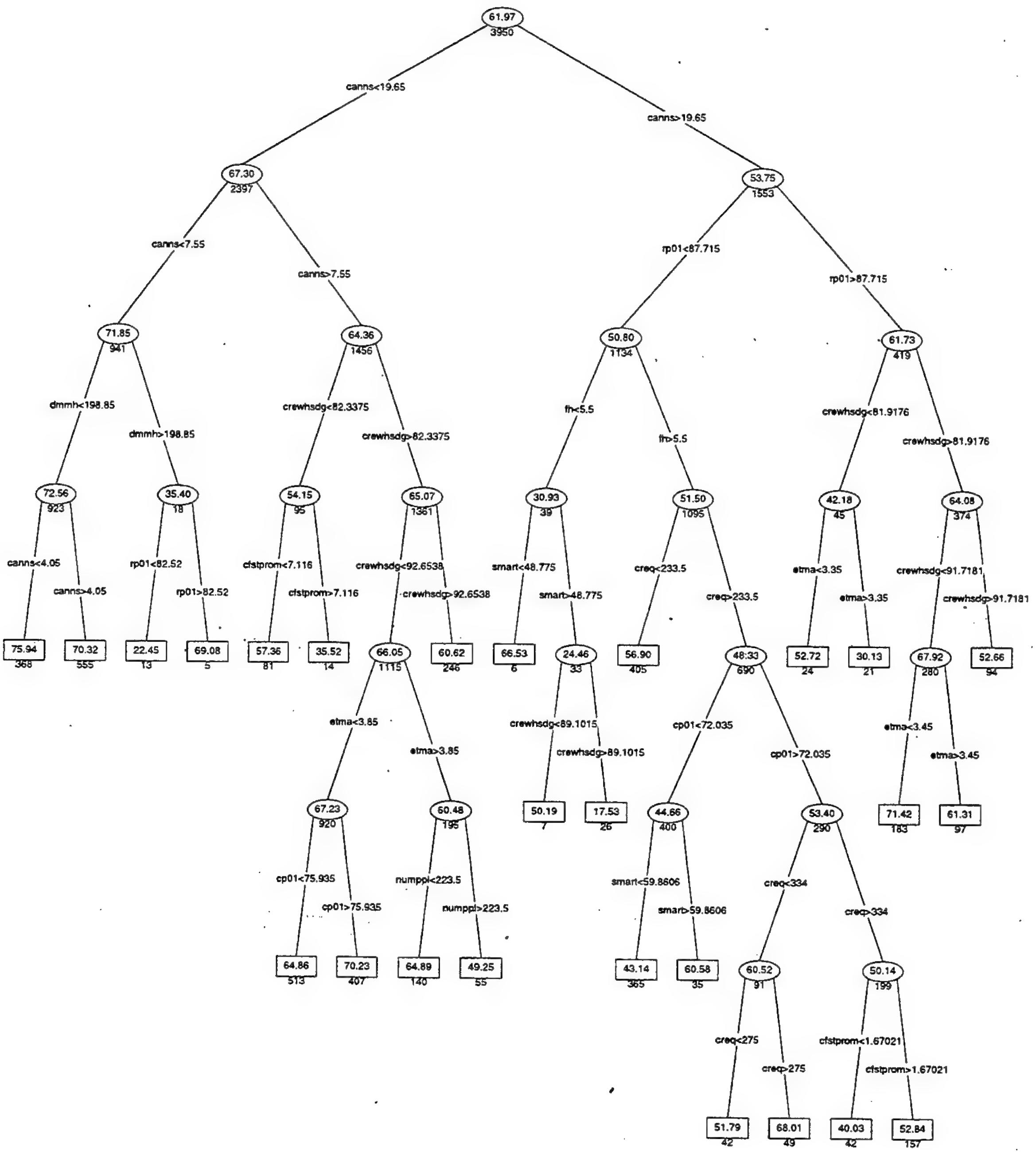


Figure 5: Tree Model Including AVCAL Variables

F. PREDICTIONS

A summary of the predicted results is depicted in Table 8. As with the data set without AVCAL variables, the test set is the 33% of the data held out for testing the predictive powers of the models. The remaining subsets were built from the original data set and may include observations used to build the models. The table lists the mean and standard deviation of the subsets' historical data as well as the predicted mean and the standard error of prediction from the full logistic regression model, the reduced logistic regression model and the regression tree model. And once again, the models with smaller standard errors predict better than taking the mean and standard deviation of the historical data with the exception of the USS George Washington during her 1996 deployment. In all cases considered, the regression tree model is the best predictor of FMC rates with the smallest standard error of prediction.

When the models are used to predict whether or not the deployment or aircraft TMS will meet CNO's FMC rate goals, (i.e., ready or not ready) the models predict well for all subsets. This may provide an additional indication that the inclusion of AVCAL variables in model development is essential in improving the predictive power of the model.

Table 8: Data with AVCAL Variables

SUBSET	MEAN	STD ERR
Test Set (33% of Data)		
Historical	62.25	18.32
Full Model	61.84	18.29
Reduced Model	61.93	18.29
Regression Tree	61.81	16.47
USS George Washington		
Historical	81.46	7.87
Full Model	65.99	7.86
Reduced Model	69.87	7.86
Regression Tree	67.82	8.15
USS Carl Vinson		
Historical	67.89	18.02
Full Model	68.18	18.00
Reduced Model	69.07	17.99
Regression Tree	64.76	14.85
USS Independence		
Historical	67.03	22.34
Full Model	62.20	22.34
Reduced Model	63.38	22.35
Regression Tree	63.64	20.13
F-14A		
Historical	60.75	18.46
Full Model	60.07	18.43
Reduced Model	60.51	18.44
Regression Tree	59.96	16.29
F/A-18C		
Historical	71.09	14.88
Full Model	65.59	14.87
Reduced Model	66.62	14.87
Regression Tree	66.77	13.80

As a final check of the goodness of fit of each model, the predictive ability of the models is tested on the remaining 1,908 observations accounting for 33% of the sample. Of these 1,908 observations, 1,067 are actually "ready". Each observation is classified in relation to its predicted probability compared with a threshold value of 61%. The naive model assumes that all aircraft TMS are ready and has an error rate of 44.09%. Table 9 provides a summary of the prediction results.

Table 9: Model Prediction Summary (AVCAL Variables)

Model	Threshold Probability	Correct Predictions (out of 1908)	Errors	Gain	% Gain over Naive Prediction
Naive		1067	841		
Full Model	0.61	1288	620	221	20.71
Reduced Model	0.61	1250	658	183	17.15
Regression Tree	**0.61**	**1309**	**599**	**242**	**22.68**

The column "errors" shows the total type I and type II errors. The regression tree model, shown in bold, provides the greatest improvement in the percent gain over the naive prediction in correct predictions with 22.68%.

V. CONCLUSIONS

A. ADDRESSING THE HYPOTHESES

1. Repairable and Consumable Items

In the logistic regression models that include AVCAL variables, FMC rates increase as the percentage of requests for consumables that are filled within 1 or 2 days increases. This is an indication of a logistics supply pipeline that is working well and that the supply personnel have the right parts on hand in sufficient quantities to prevent off-ship ordering of essential parts. The same models indicate that as the number of requests for consumable items increases, FMC rates decrease. In addition to the two variables listed above, the regression tree also identifies the percentage of requests for repairable items that were filled in 1 or 2 days as a significant variable in predicting FMC rates.

2. Cannibalization Rates

An increase in cannibalizations alone decreases FMC rates, while an increase in the number of sorties causes FMC rates to increase. CNA also came to the conclusion that an increase in the number of sorties increases the FMC rate, which supported their assumption that machinery needs to be used to keep it running well [Ref. 9]. However, all models identify that when considered together, an increase in the number of cannibalizations combined with an increase in the number of sorties flown by the squadron have the most significant negative effect on FMC rates. Cannibalization is a symptom of a supply system that does not have the right parts on hand. The problem is magnified as the number of sorties increases and maintenance personnel are forced to cannibalize for necessary parts to meet operational commitments.

3. Flight Hours

Flight hours were not identified as significant predictors in the logistic regression models but were found to be significant in the regression trees.

4. The Quality and Quantity of Personnel

The percentage of the crew with high school degrees, the percentage of crew who scored in the upper mental group on the AFQT, the number of enlisted people assigned to the squadron, and the percentage of the crew attached to the squadron compared with M+1 requirements weighted by paygrade were considered significant variables in predicting FMC rates. Not all interactions were easily explained. In the logistic regression model as the percentage of the crew with high school degrees increases and the percentage who score in the upper mental group on the AFQT increases, FMC rates decrease. There is also an intangible difference when analyzing the statistics on the quantity and quality of the personnel. The USS George Washington completed her 1996 deployment with an FMC rate of over 80%. The models that predicted fairly accurately for the other carriers did not predict well for her. The end-of-deployment lessons learned report hints at a group of people who aggressively and proactively addressed potential logistic pipeline concerns during both the pre-deployment and the deployment phases. The individual drive to excel is hard to anticipate or predict.

5. Model Flexibility

In general, the models predict well for a deployed carrier or a specific aircraft TMS. However, as indicated in Tables 5 and 9, the models predict accurately for a single observation about 65% of the time. The regression tree with AVCAL variables is the best model for predicting a single observation, predicting correctly 68.61% of the time.

B. ADDRESSING THE RESEARCH QUESTIONS

1. Significant Predictive Factors

This research indicates that several factors are significant in predicting FMC rates. From the logistic regression models, the interaction between cannibalizations and sorties as each increases has the greatest negative impact on FMC rates. Other predictors which significantly influenced FMC rates are the percentage of items which can not be repaired at the AIMD and have to be sent off the ship for depot-level maintenance, elapsed maintenance time per maintenance action, the percentage of requests for consumable items that are filled in 1 or 2 days, and the number of requests for consumable items.

The regression trees also indicate a similar list of significant predictors with cannibalizations playing a prominent role in building the trees. However, in the tree models there is a greater emphasis on personnel quality and quantity factors. The regression trees include the percentage of the crew with a high school degree, the percentage of the crew who score in the upper mental group on the AFQT, and percentage of crew who were not with the squadron six months earlier.

Several of the findings compare well with CNA's results, which included enlisted personnel quality, enlisted turnover, deployed sorties, and supplies on hand as significant drivers of the FMC rate [Ref. 9].

2. Comparing Logistic Regression and Regression Tree Methodologies

Comparing and contrasting the two methodologies is the second research question. Because the response is binary, both logistic regression and regression trees are well-suited for the problem. The models produced consistent probabilistic outcomes that center on FMC rates. The regression tree did reveal structure within the data while

ignoring variable interactions. The tree models provided a better fit for FMC rate analysis.

3. Implications of Resulting Predictive Models

There are three main implications of these predictive models. The first is that the models can predict FMC rates with some success and are an improvement over having no model at all. The second implication is that AVCAL variables are significant factors in predicting FMC rates. Both the large logistic regression and regression tree models with AVCAL variables predicted more accurately than those excluding AVCAL variables. The analysis would have been much more useful with one complete data set, which included all aircraft TMS, as well as AVCAL variables. The third implication is that in addition to the models' predictive power, resource managers are provided with a list of significant predictive factors on which to focus time and money in an effort to improve aircraft readiness.

LIST OF REFERENCES

1. Ackart, Leigh P. *An Evaluation of Markov Chain Modeling for F/A-18 Aircraft Readiness,* M.S. Thesis in Operations Research, Naval Postgraduate School, September 1998.

2. Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone. (1984) *Classification and Regression Trees.* Waysworth, Inc.

3. Chambers, John M. and Trevor J. Hastie. (1992) *Statistical Models in S.* AT&T Bell Laboratories.

4. Collett D.(1991) *Modelling Binary Data.* London: Chapman and Hall.

5. Executive briefing for the Chief of Naval Operations, *Review of Material Support for Naval Aviation,* November 1997.

6. Francis, Peter and Jessica S. Oi. Center for Naval Analyses, *Volatility in Readiness Measures,* CRM 97-128, January 1998.

7. Hamilton, Lawrence C. (1992) *Regression with Graphics.* Wadsworth, Inc.

8. Joint Publication 1-02, Department of Defense Dictionary of Military and Associated Terms, As Amended 10 June 1998.

9. Junor, Laura J., James M. Jondrow, Peter Francis, and Jessica S. Oi. Center for Naval Analyses, *Understanding Aircraft Readiness: An Empirical Approach,* CRM 97-014, March 1997.

10. Junor, Laura J., James M. Jondrow, Peter Francis, Robert P. Trost, and Jessica S. Oi. Center for Naval Analyses, *Managing Readiness,* CRM 97-127, January 1998.

11. Office of the Chief of Naval Operations letter 4441 N881E1/3U657111 dated 19 January 1994.

12. Pennypacker, Blaine S. *A Comparison and Validation of Two Surface Ship Readiness Models,* Naval Postgraduate School, September 1994.

13. Ryan, Thomas P. (1997) *Modern Regression Methods.* John Wiley and Sons, Inc.

14. S-Plus, *Guide to Statistical & Mathematical Analysis,* MathSoft, Inc., 1988,1995.

15. USS George Washington, "Post-Deployment Report," 1996. Photocopied.

16. Van Brabant, John. *A Monthly Squadron Sortie Scheduling Model for Improved Combat Readiness,* M.S. Thesis in Operations Research, Naval Postgraduate School, September 1993.

17. Venables, W.N., and B.D. Ripley. (1994) *Modern Applied Statistics with S-Plus.* Springer-Verlag New York, Inc.

18. Will, Jonathan E. *Modelling DD-963 Class Material Readiness,* Naval Postgraduate School, September 1988.

Appendix. Tree Model Output

This appendix contains the S-Plus tree outputs for the data set without AVCAL variables and the data set with AVCAL variables. Each row contains the node split, the number of cases in the node, the deviance at the node, and the predicted FMC rate at that node. A row ending with a "*" indicates a terminal node.

1. Tree without AVCAL variables, 23 terminal nodes

```
1) root 7361 3106000.0 60.31
  2) canns<15.15 3486 1217000.0 66.51
    4) canns<7.55 1479  527800.0 69.85
      8) sorties<31 108  111600.0 55.44
        16) turn6mo<25.2788 89   81490.0 61.48 *
        17) turn6mo>25.2788 19   11650.0 27.17 *
      9) sorties>31 1371  392000.0 70.98
        18) bcm<56.55 1347  357600.0 71.38 *
        19) bcm>56.55 24   21890.0 48.37 *
    5) canns>7.55 2007  660900.0 64.05
      10) bcm<30.75 1070  310700.0 66.62
        20) crewhsdg<81.6543 65   20710.0 53.32 *
        21) crewhsdg>81.6543 1005  277800.0 67.48 *
      11) bcm>30.75 937  335000.0 61.12
        22) sorties<80.5 286  136400.0 55.95 *
        23) sorties>80.5 651  187700.0 63.38 *
  3) canns>15.15 3875 1634000.0 54.73
    6) crewhsdg<81.587 391  139700.0 44.23 *
    7) crewhsdg>81.587 3484 1446000.0 55.91
      14) bcm<34.65 2607  980100.0 57.78
        28) canns<24.95 1166  416000.0 61.41
          56) etma<4.45 1055  356700.0 62.39
            112) crewhsdg<91.0927 733  233500.0 64.52 *
            113) crewhsdg>91.0927 322  112300.0 57.53 *
          57) etma>4.45 111   48840.0 52.16 *
        29) canns>24.95 1441  536200.0 54.84
          58) crewhsdg<91.8721 1131  411200.0 56.23
            116) smart<51.1793 323  104900.0 51.32 *
            117) smart>51.1793 808  295400.0 58.19 *
          59) crewhsdg>91.8721 310  115000.0 49.78
            118) fh<2.5 5     354.7  7.18 *
            119) fh>2.5 305  105400.0 50.48 *
      15) bcm>34.65 877  430200.0 50.36
        30) fh<2.5 14    5780.0 15.24 *
```

```
31) fh>2.5 863  406800.0 50.93
  62) smart<52.4088 192  67940.0 58.79 *
  63) smart>52.4088 671  323600.0 48.68
   126) bcm<51.8 622  275400.0 49.75
     252) dmmh<162.45 561  228600.0 48.55
       504) canns<30.75 406  162900.0 51.18 *
       505) canns>30.75 155  55520.0 41.65 *
     253) dmmh>162.45 61  38550.0 60.80 *
   127) bcm>51.8 49  38570.0 35.15
     254) smart<61.2614 31  18460.0 47.50 *
     255) smart>61.2614 18  7236.0 13.88 *
```

2. Tree with AVCAL variables, 26 terminal nodes

```
1) root 3950 1406000 61.97
  2) canns<19.65 2397  645800 67.30
    4) canns<7.55 941  232600 71.85
      8) dmmh<198.85 923  187100 72.56
        16) canns<4.05 368  66990 75.94 *
        17) canns>4.05 555  113200 70.32 *
      9) dmmh>198.85 18  21100 35.40
        18) rp01<82.52 13  8806 22.45 *
        19) rp01>82.52 5  4441 69.08 *
    5) canns>7.55 1456  381100 64.36
      10) crewhsdg<82.3375 95  30660 54.15
        20) cfstprom<7.116 81  21120 57.36 *
        21) cfstprom>7.116 14  3849 35.52 *
      11) crewhsdg>82.3375 1361  339900 65.07
        22) crewhsdg<92.6538 1115  271100 66.05
          44) etma<3.85 920  200900 67.23
            88) cp01<75.935 513  109300 64.86 *
            89) cp01>75.935 407  85010 70.23 *
          45) etma>3.85 195  62790 60.48
            90) numppl<223.5 140  38920 64.89 *
            91) numppl>223.5 55  14210 49.25 *
        23) crewhsdg>92.6538 246  62870 60.62 *
  3) canns>19.65 1553  587400 53.75
    6) rp01<87.715 1134  411700 50.80
      12) fh<5.5 39  27820 30.93
        24) smart<48.775 6  2271 66.53 *
        25) smart>48.775 33  16560 24.46
          50) crewhsdg<89.1015 7  2780 50.19 *
          51) crewhsdg>89.1015 26  7902 17.53 *
      13) fh>5.5 1095  368000 51.50
        26) creq<233.5 405  118400 56.90 *
        27) creq>233.5 690  230800 48.33
```

```
    54) cp01<72.035 400  117900 44.66
     108) smart<59.8606 365   98220 43.14 *
     109) smart>59.8606 35    9998 60.58 *
    55) cp01>72.035 290  100000 53.40
     110) creq<334 91   33260 60.52
       220) creq<275 42   17540 51.79 *
       221) creq>275 49    9772 68.01 *
     111) creq>334 199   60040 50.14
       222) cfstprom<1.67021 42    9484 40.03 *
       223) cfstprom>1.67021 157   45120 52.84 *
7) rp01>87.715 419  139100 61.73
 14) crewhsdg<81.9176 45  15840 42.18
   28) etma<3.35 24    5681 52.72 *
   29) etma>3.35 21    4443 30.13 *
 15) crewhsdg>81.9176 374  104000 64.08
   30) crewhsdg<91.7181 280   66680 67.92
     60) etma<3.45 183   36700 71.42 *
     61) etma>3.45 97   23500 61.31 *
   31) crewhsdg>91.7181 94   20940 52.66 *
```

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, Virginia 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, California 93943-5101

3. Defense Logistic Studies Information Exchange 1
   U.S. Army Logistics Management Center
   Fort Lee, Virginia 23801-6043

4. Deputy Chief of Naval Operations (Logistics) 1
   ATTN: LCDR Carolyn Kresek, N421C
   2000 Navy Pentagon
   Washington, DC 20350-2000

5. Professor Sam Buttrey, Code OR/Sb 1
   1411 Cunningham Rd.
   Naval Postgraduate School
   Monterey, California 93943

6. Professor Lyn Whitaker, Code OR/Lw 1
   1411 Cunningham Rd.
   Naval Postgraduate School
   Monterey, California 93943

7. Patricia B. Moore, LCDR, USN 2
   Atlantic Fleet Headquarters Support Activity
   7918 Blandy Road Suite 100
   Norfolk, Virginia 23551-2419

8. Dr. David Schrady 1
   Code OR/So
   Naval Postgraduate School
   Monterey, California 93943-5000